Claims:

1. A method of restructuring an input HTML document to comply with strict HTML, said method comprising the steps of:

(a) linearly traversing said input HTML document to create a hierarchical tree structure representation, said traversal maintaining a current insertion point for elements within said tree structure representation;

(b) during said traversal, identifying those elements of said input HTML document that violate strict HTML and, for each said identified element:

(b)(i) retracing said tree structure representation from said current insertion point to identify an further insertion point from which said identified element can depend, said retracing comprising noting each parent element of said identified element passed during said retracing;

(b)(ii) at said further insertion point, creating new elements in said tree structure representation corresponding to those said parent elements passed during said retracing, said new elements being created in reverse chronological order to that encountered during said retracing; and

(b)(iii) appending said identified element to a terminal one of said new elements; and

(c) converting said tree structure representation into an output HTML document.

2. A method according to claim 1, wherein step (b)(iii) comprises creating a link from said appended identified element to a first said parent element encountered during said retracting.

3. A method according to claim 2 wherein said link comprises a vector.

4. A method according to claim 1 wherein step (b)(iii) comprises copying a syntax of a first said parent element encountered during said retracting to said appended identified element.

5. A method according to claim 1, wherein one or more of said elements comprises information associated therewith, said method comprising the further steps, before step (a) of:

performing an initial pass of said input HTML document to identify said elements having said associated information, and maintaining a record of each such element and the corresponding associated information whereby each time said element is placed in said tree structure representation, the corresponding associated information is associated therewith.

6. A method according to claim 1 comprising the further step of:

(d) reproducing said output HTML document.

7. A method according to claim 6 wherein said output HTML document is reproduced using a video display.

8. A method according to claim 6 wherein said output HTML document is reproduced using a printer.

9. Apparatus for restructuring an input HTML document to comply with strict HTML, said apparatus comprising:

means for linearly traversing said input HTML document to create a hierarchical tree structure representation, said traversal maintaining a current insertion point for elements within said tree structure representation;

means for identifying, during said traversal, an element of said input HTML document that violates strict HTML;

means for retracing said tree structure representation from said current insertion point to identify an further insertion point from which said identified element can depend, said retracing comprising noting each parent element of said identified element passed during said retracing;

means for creating, at said further insertion point, at leats one new elemcnt in said tree structure representation corresponding to those said parent elements passed during said retracing, said new elements being created in reverse chronological order to that encountered during said retracing; and

means for appending said identified element to a terminal one of said new elements; and

means for converting said tree structurc representation into an output HTML document.

10. Apparatus according to claim 9, wherein said means for appending creates a link from said appended identified element to a first said parent element encountered during said retracting.

11. Apparatus according to claim 10 wherein said link comprises a vector.

12. Apparatus according to claim 9 wherein said means for appending copies a syntax of a first said parent element encountered during said retracting to said appended identified element.

13. Apparatus according to claim 9, wherein one or more of said elements comprises information associated therewith, and said apparatus further comprises:

means for performing an initial pass of said input HTML document to identify said elements having said associated information, and for maintaining a record of each such element and the corresponding associated information whereby each time said element is placed in said tree structure representation, the corresponding associated information is associated therewith.

14. Apparatus according to claim 9 further comprising:

means for reproducing said output HTML document.

15. Apparatus according to claim 14 wherein said means for reproducing comprises a video display.

16. Apparatus according to claim 14 wherein said means for reproducing comprises a printer.

17. A computer readable medium, having a program recorded thereon, where the program is configured to make a computer execute a procedure to restructure an input HTML document to comply with strict HTML, said program comprising:

code for linearly traversing said input HTML document to create a hierarchical tree structure representation, said traversal maintaining a current insertion point for elements within said tree structure representation;

code for, during said traversal, identifying an element of said input HTML document that violate strict HTML ;

code for retracing said tree structure representation from said current insertion point to identify an further insertion point from which said identified element can depend, said retracing comprising noting each parent element of said identified element passed during said retracing;

code for creating new elements in said tree structure representation at said further insertion point, said new elements corresponding to those said parent elements passed during said retracing, said new elements being created in reverse chronological order to that encountered during said retracing; and

code for appending said identified element to a terminal one of said new elements; and

code for converting said tree structure representation into an output HTML document.

18. A computer readable medium according to claim 17, wherein code for appending comprises code for creating a link from said appended identified element to a first said parent element encountered during said retracting.

19. A computer readable medium according to claim 18 wherein said link comprises a vector.

20. A computer readable medium according to claim 17 wherein said code for appending comprises code for copying a syntax of a first said parent element encountered during said retracting to said appended identified element.

21. A computer readable medium according to claim 17, wherein one or more of said elements comprises information associated therewith, and said program further comprises:

code for performing an initial pass of said input HTML document to identify said elements having said associated information, and for maintaining a record of each such element and the corresponding associated information whereby each time said element is placed in said tree structure representation, the corresponding associated information is associated therewith.

22. A computer readable medium according to claim 17, said program further comprising:

code for reproducing said output HTML document.

23. A computer readable medium according to claim 22 wherein said output HTML document is reproduced using a video display.

24. A computer readable medium according to claim 22 wherein said output HTML document is reproduced using a printer.